

#8/4B

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: KAJIWARA, Susumu
MISAWA, Norihiko
KONDO, Keiji
- (ii) TITLE OF INVENTION: A DNA CHAIN USEFUL FOR INCREASING
PRODUCTION OF CAROTENOIDS
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: FOLEY & LARDNER
 - (B) STREET: 3000 K Street, N.W.
 - (C) CITY: Washington
 - (D) STATE: D.C.
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 20007-5109
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/737,319
 - (B) FILING DATE: 12-NOV-1996
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT/JP96/00574
 - (B) FILING DATE: 08-MAR-1996
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: JP 51234/1995
 - (B) FILING DATE: 10-MAR-1995
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Bent, Stephen A.
 - (B) REGISTRATION NUMBER: 29,768
 - (C) REFERENCE/DOCKET NUMBER: 081356/0111
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (202) 672-5300
 - (B) TELEFAX: (202) 672-5399

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 251 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Ser	Met	Pro	Asn	Ile	Val	Pro	Pro	Ala	Glu	Val	Arg	Thr	Glu	Gly	1		5		10		15							
Leu	Ser	Leu	Glu	Glu	Tyr	Asp	Glu	Glu	Gln	Val	Arg	Leu	Met	Glu	Glu		20		25		30								
Arg	Cys	Ile	Leu	Val	Asn	Pro	Asp	Asp	Val	Ala	Tyr	Gly	Glu	Ala	Ser		35		40		45								
Lys	Lys	Thr	Cys	His	Leu	Met	Ser	Asn	Ile	Asn	Ala	Pro	Lys	Asp	Leu		50		55		60								
Leu	His	Arg	Ala	Phe	Ser	Val	Phe	Leu	Phe	Arg	Pro	Ser	Asp	Gly	Ala		65		70		75								
Leu	Leu	Leu	Gln	Arg	Arg	Ala	Asp	Glu	Lys	Ile	Thr	Phe	Pro	Gly	Met			85		90									
Trp	Thr	Asn	Thr	Cys	Cys	Ser	His	Pro	Leu	Ser	Ile	Lys	Gly	Glu	Val			100		105									
Glu	Glu	Glu	Asn	Gln	Ile	Gly	Val	Arg	Arg	Ala	Ala	Ser	Arg	Lys	Leu			115		120									
Glu	His	Glu	Leu	Gly	Val	Pro	Thr	Ser	Ser	Thr	Pro	Pro	Asp	Ser	Phe			130		135									
Thr	Tyr	Leu	Thr	Arg	Ile	His	Tyr	Leu	Ala	Pro	Ser	Asp	Gly	Leu	Trp				150		155								
Gly	Glu	His	Glu	Ile	Asp	Tyr	Ile	Leu	Phe	Ser	Thr	Thr	Pro	Thr	Glu				165		170								
His	Thr	Gly	Asn	Pro	Asn	Glu	Val	Ser	Asp	Thr	Arg	Tyr	Val	Thr	Lys				180		185								
Pro	Glu	Leu	Gln	Ala	Met	Phe	Glu	Asp	Glu	Ser	Asn	Ser	Phe	Thr	Pro				195		200								
Trp	Phe	Lys	Leu	Ile	Ala	Arg	Asp	Phe	Leu	Phe	Gly	Trp	Trp	Asp	Gln					215		220							
Leu	Leu	Ala	Arg	Arg	Asn	Glu	Lys	Gly	Glu	Val	Asp	Ala	Lys	Ser	Leu					230		235							
Glu	Asp	Leu	Ser	Asp	Asn	Lys	Val	Trp	Lys	Met																			

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 259 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gln Leu Leu Ala Glu Asp Arg Thr Asp His Met Arg Gly Ala Ser
1 5 10 15
Thr Trp Ala Gly Gly Gln Ser Gln Asp Glu Leu Met Leu Lys Asp Glu
20 25 30
Cys Ile Leu Val Asp Ala Asp Asp Asn Ile Thr Gly His Val Ser Lys
35 40 45
Leu Glu Cys His Lys Phe Leu Pro His Gln Pro Ala Gly Leu Leu His
50 55 60
Arg Ala Phe Ser Val Phe Leu Phe Asp Asp Gln Gly Arg Leu Leu Leu
65 70 75 80
Gln Gln Arg Ala Arg Ser Lys Ile Thr Phe Pro Ser Val Trp Thr Asn
85 90 95
Thr Cys Cys Ser His Pro Leu His Gly Gln Thr Pro Asp Glu Val Asp
100 105 110
Gln Leu Ser Gln Val Ala Asp Gly Thr Val Pro Gly Ala Lys Ala Ala
115 120 125
Ala Ile Arg Lys Leu Glu His Glu Leu Gly Ile Pro Ala His Gln Leu
130 135 140
Pro Ala Ser Ala Phe Arg Phe Leu Thr Arg Leu His Tyr Cys Ala Ala
145 150 155 160
Asp Val Gln Pro Ala Ala Thr Gln Ser Ala Leu Trp Gly Glu His Glu
165 170 175
Met Asp Tyr Ile Leu Phe Ile Arg Ala Asn Val Thr Leu Ala Pro Asn
180 185 190
Pro Asp Glu Val Asp Glu Val Arg Tyr Val Thr Gln Glu Glu Leu Arg
195 200 205
Gln Met Met Gln Pro Asp Asn Gly Leu Gln Trp Ser Pro Trp Phe Arg
210 215 220
Ile Ile Ala Ala Arg Phe Leu Glu Arg Trp Trp Ala Asp Leu Asp Ala
225 230 235 240
Ala Leu Asn Thr Asp Lys His Glu Asp Trp Gly Thr Val His His Ile
245 250 255
Asn Glu Ala

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Thr	Ala	Asp	Asn	Asn	Ser	Met	Pro	His	Gly	Ala	Val	Ser	Ser	Tyr	
1				5					10					15		
Ala	Lys	Leu	Val	Gln	Asn	Gln	Thr	Pro	Glu	Asp	Ile	Leu	Glu	Glu	Phe	
			20					25					30			
Pro	Glu	Ile	Ile	Pro	Leu	Gln	Gln	Arg	Pro	Asn	Thr	Arg	Ser	Ser	Glu	
		35					40					45				
Thr	Ser	Asn	Asp	Glu	Ser	Gly	Glu	Thr	Cys	Phe	Ser	Gly	His	Asp	Glu	
	50					55					60					
Glu	Gln	Ile	Lys	Leu	Met	Asn	Glu	Asn	Cys	Ile	Val	Leu	Asp	Trp	Asp	
65					70					75					80	
Asp	Asn	Ala	Ile	Gly	Ala	Gly	Thr	Lys	Lys	Val	Cys	His	Leu	Met	Glu	
				85					90					95		
Asn	Ile	Glu	Lys	Gly	Leu	Leu	His	Arg	Ala	Phe	Ser	Val	Phe	Ile	Phe	
		100						105					110			
Asn	Glu	Gln	Gly	Glu	Leu	Leu	Leu	Gln	Gln	Arg	Ala	Thr	Glu	Lys	Ile	
		115					120						125			
Thr	Phe	Pro	Asp	Leu	Trp	Thr	Asn	Thr	Cys	Cys	Ser	His	Pro	Leu	Cys	
	130					135					140					
Ile	Asp	Asp	Glu	Leu	Gly	Leu	Lys	Gly	Lys	Leu	Asp	Asp	Lys	Ile	Lys	
145					150					155					160	
Gly	Ala	Ile	Thr	Ala	Ala	Val	Arg	Lys	Leu	Asp	His	Glu	Leu	Gly	Ile	
				165					170					175		
Pro	Glu	Asp	Glu	Thr	Lys	Thr	Arg	Gly	Lys	Phe	His	Phe	Leu	Asn	Arg	
		180						185					190			
Ile	His	Tyr	Met	Ala	Pro	Ser	Asn	Glu	Pro	Trp	Gly	Glu	His	Glu	Ile	
		195					200					205				
Asp	Tyr	Ile	Leu	Phe	Tyr	Lys	Ile	Asn	Ala	Lys	Glu	Asn	Leu	Thr	Val	
	210					215					220					
Asn	Pro	Asn	Val	Asn	Glu	Val	Arg	Asp	Phe	Lys	Trp	Val	Ser	Pro	Asn	
225				230						235					240	
Asp	Leu	Lys	Thr	Met	Phe	Ala	Asp	Pro	Ser	Tyr	Lys	Phe	Thr	Pro	Trp	
				245					250					255		
Phe	Lys	Ile	Ile	Cys	Glu	Asn	Tyr	Leu	Phe	Asn	Trp	Trp	Glu	Gln	Leu	
		260						265					270			
Asp	Asp	Leu	Ser	Glu	Val	Glu	Asn	Asp	Arg	Gln	Ile	His	Arg	Met	Leu	
		275					280					285				

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1099 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 99..851

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCCACGCGTC CGCACATCTC GCATATATCA CTTTCCTCCT TCCAGAACAA GTTCTGAGTC	60
AACCGAAAAG AAAGAAGGCA GAGGAAAATA TATTCTAG ATG TCC ATG CCC AAC	113
Met Ser Met Pro Asn	
1 5	
ATT GTT CCC CCC GCC GAG GTC CGA ACC GAA GGA CTC AGT TTA GAA GAG	161
Ile Val Pro Pro Ala Glu Val Arg Thr Glu Gly Leu Ser Leu Glu Glu	
10 15 20	
TAC GAT GAG GAG CAG GTC AGG CTG ATG GAG GAG CGA TGT ATT CTT GTT	209
Tyr Asp Glu Glu Gln Val Arg Leu Met Glu Glu Arg Cys Ile Leu Val	
25 30 35	
AAC CCG GAC GAT GTG GCC TAT GGA GAG GCT TCG AAA AAG ACC TGC CAC	257
Asn Pro Asp Asp Val Ala Tyr Gly Glu Ala Ser Lys Lys Thr Cys His	
40 45 50	
TTG ATG TCC AAC ATC AAC GCG CCC AAG GAC CTC CTC CAC CGA GCA TTC	305
Leu Met Ser Asn Ile Asn Ala Pro Lys Asp Leu Leu His Arg Ala Phe	
55 60 65	
TCC GTG TTT CTC TTC CGC CCA TCG GAC GGA GCA CTC CTG CTT CAG CGA	353
Ser Val Phe Leu Phe Arg Pro Ser Asp Gly Ala Leu Leu Leu Gln Arg	
70 75 80 85	
AGA GCG GAC GAG AAG ATT ACG TTC CCT GGA ATG TGG ACC AAC ACG TGT	401
Arg Ala Asp Glu Lys Ile Thr Phe Pro Gly Met Trp Thr Asn Thr Cys	
90 95 100	
TGC AGT CAT CCT TTG AGC ATC AAG GGC GAG GTT GAA GAG GAG AAC CAG	449
Cys Ser His Pro Leu Ser Ile Lys Gly Glu Val Glu Glu Glu Asn Gln	
105 110 115	
ATC GGT GTT CGA CGA GCT GCG TCC CGA AAG TTG GAG CAC GAG CTT GGC	497
Ile Gly Val Arg Arg Ala Ala Ser Arg Lys Leu Glu His Glu Leu Gly	
120 125 130	
GTG CCT ACA TCG TCG ACT CCG CCC GAC TCG TTC ACC TAC CTC ACT AGG	545
Val Pro Thr Ser Ser Thr Pro Pro Asp Ser Phe Thr Tyr Leu Thr Arg	
135 140 145	
ATA CAT TAC CTC GCT CCG AGT GAC GGA CTC TGG GGA GAA CAC GAG ATC	593
Ile His Tyr Leu Ala Pro Ser Asp Gly Leu Trp Gly Glu His Glu Ile	
150 155 160 165	

GAC TAC ATT CTC TTC TCA ACC ACA CCT ACA GAA CAC ACT GGA AAC CCT	641
Asp Tyr Ile Leu Phe Ser Thr Thr Pro Thr Glu His Thr Gly Asn Pro	
170 175 180	
AAC GAA GTC TCT GAC ACT CGA TAT GTC ACC AAG CCC GAG CTC CAG GCG	689
Asn Glu Val Ser Asp Thr Arg Tyr Val Thr Lys Pro Glu Leu Gln Ala	
185 190 195	
ATG TTT GAG GAC GAG TCT AAC TCA TTT ACC CCT TGG TTC AAG TTG ATT	737
Met Phe Glu Asp Glu Ser Asn Ser Phe Thr Pro Trp Phe Lys Leu Ile	
200 205 210	
GCC CGA GAC TTC CTG TTT GGC TGG TGG GAT CAA CTT CTC GCC AGA CGA	785
Ala Arg Asp Phe Leu Phe Gly Trp Trp Asp Gln Leu Leu Ala Arg Arg	
215 220 225	
AAT GAA AAG GGT GAG GTC GAT GCC AAA TCG TTG GAG GAT CTC TCG GAC	833
Asn Glu Lys Gly Glu Val Asp Ala Lys Ser Leu Glu Asp Leu Ser Asp	
230 235 240 245	
AAC AAA GTC TGG AAG ATG TAGTCGACCC TTCTTTCTGT ACAGTCATCT	881
Asn Lys Val Trp Lys Met	
250	
CAGTTGCGCCT GTTGGTTGCT TGCTTCTTGC TCTTCTTTCT ATATATCTTT TTTCTTGCTC	941
GGGTAGACTT GATCTTTCTA CATAGCATAC GCATACATAC ATAAACTCTA TTTCTTGCTC	1001
TTTATCTCTC TTCTAAGGGA ATCTTCAAGA TCAATTTCTT TTTGGGCTAC AACATTTTCAG	1061
ATCAATGTTG CTTTTTCAGAC TACAAAAAAA AAAAAAAA	1099

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1074 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:
- (A) NAME/KEY: CDS
 - (B) LOCATION: 145..921

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATCGCTACTT GGAACCTGGC CCGGCGGCAG TCCGATGACG CGATGCTTCG TTCGTTGCTC	60
AGAGGCCTCA CGCATTTCCC CCGCGTGAAC TCCGCGCAGC AGCCCAGCTG TGCACACGCG	120
CGACTCCAGT TTAGGCCCAG AAGC ATG CAG CTG CTT GCC GAG GAC CGC ACA	171
Met Gln Leu Leu Ala Glu Asp Arg Thr	
1 5	
GAC CAT ATG AGG GGT GCA AGT ACC TGG GCA GGC GGG CAG TCG CAG GAT	219
Asp His Met Arg Gly Ala Ser Thr Trp Ala Gly Gly Gln Ser Gln Asp	
10 15 20 25	

GAG CTG ATG CTG AAG GAC GAG TGC ATC TTG GTG GAT GCT GAC GAC AAC	267
Glu Leu Met Leu Lys Asp Glu Cys Ile Leu Val Asp Ala Asp Asp Asn	
30 35 40	
ATT ACA GGC CAT GTC AGC AAG CTG GAG TGC CAC AAG TTC CTA CCA CAT	315
Ile Thr Gly His Val Ser Lys Leu Glu Cys His Lys Phe Leu Pro His	
45 50 55	
CAG CCT GCA GGC CTG CTG CAC CGG GCC TTC TCT GTA TTC CTG TTT GAC	363
Gln Pro Ala Gly Leu Leu His Arg Ala Phe Ser Val Phe Leu Phe Asp	
60 65 70	
GAC CAG GGG CGA CTG CTG CTG CAA CAG CGT GCA CGA TCA AAA ATC ACA	411
Asp Gln Gly Arg Leu Leu Leu Gln Gln Arg Ala Arg Ser Lys Ile Thr	
75 80 85	
TTC CCC AGT GTG TGG ACC AAC ACC TGC TGC AGC CAC CCT CTA CAT GGG	459
Phe Pro Ser Val Trp Thr Asn Thr Cys Cys Ser His Pro Leu His Gly	
90 95 100 105	
CAG ACC CCA GAT GAG GTG GAC CAA CTA AGC CAG GTG GCC GAC GGC ACA	507
Gln Thr Pro Asp Glu Val Asp Gln Leu Ser Gln Val Ala Asp Gly Thr	
110 115 120	
GTA CCT GGC GCA AAG GCT GCT GCC ATC CGC AAG TTG GAG CAC GAG CTG	555
Val Pro Gly Ala Lys Ala Ala Ala Ile Arg Lys Leu Glu His Glu Leu	
125 130 135	
GGG ATA CCA GCG CAC CAG CTG CCG GCC AGC GCG TTT CGC TTC CTC ACG	603
Gly Ile Pro Ala His Gln Leu Pro Ala Ser Ala Phe Arg Phe Leu Thr	
140 145 150	
CGT TTG CAC TAC TGC GCC GCG GAC GTG CAG CCG GCT GCG ACA CAA TCA	651
Arg Leu His Tyr Cys Ala Ala Asp Val Gln Pro Ala Ala Thr Gln Ser	
155 160 165	
GCA CTC TGG GGC GAG CAC GAA ATG GAC TAC ATC TTA TTC ATC CGG GCC	699
Ala Leu Trp Gly Glu His Glu Met Asp Tyr Ile Leu Phe Ile Arg Ala	
170 175 180 185	
AAC GTC ACC CTT GCG CCC AAC CCT GAC GAG GTG GAC GAA GTC AGG TAC	747
Asn Val Thr Leu Ala Pro Asn Pro Asp Glu Val Asp Glu Val Arg Tyr	
190 195 200	
GTG ACG CAG GAG GAG CTG CGG CAG ATG ATG CAG CCG GAC AAT GGG TTG	795
Val Thr Gln Glu Glu Leu Arg Gln Met Met Gln Pro Asp Asn Gly Leu	
205 210 215	
CAA TGG TCG CCG TGG TTT CGC ATC ATC GCC GCG CGC TTC CTT GAG CGC	843
Gln Trp Ser Pro Trp Phe Arg Ile Ile Ala Ala Arg Phe Leu Glu Arg	
220 225 230	
TGG TGG GCT GAC CTA GAC GCG GCC CTG AAC ACT GAC AAA CAC GAG GAT	891
Trp Trp Ala Asp Leu Asp Ala Ala Leu Asn Thr Asp Lys His Glu Asp	
235 240 245	
TGG GGA ACG GTG CAT CAC ATC AAC GAA GCG TGAAAACAGA AGCTGTAGGA	941
Trp Gly Thr Val His Ile Asn Glu Ala	
250 255	
TGTCAAGACA CGTCATGAGG GGGCTTGCCA TCTTGCGCGC TTCGTATCTC TTTTACTGA	1001

GACTGAACCT GCAGCTGGAG ACAATGGTGA GCCCAATTCA ACTTTCCGCT GCACTGGAAA 1061
 AAAAAAAAAA AAA 1074

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1058 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 187..1050

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TCGATGGGGG TTGCCTTTCT TTTTCGGTCT TAACTCCATT TATATTTATT TATTCATTTT	60
TATCTATTTA ACAGGAAACA GTTTTCTAGT GACAAGAAGG CGTATATCCC ACTTAATTCA	120
ATATTAGAGT ATTCGTATTT GGAATACAGG AAGAGTAAAA ATAAGCCAAA AATTCATTAC	180
ACCTCA ATG ACT GCC GAC AAC AAT AGT ATG CCC CAT GGT GCA GTA TCT	228
Met Thr Ala Asp Asn Asn Ser Met Pro His Gly Ala Val Ser	
1 5 10	
AGT TAC GCC AAA TTA GTG CAA AAC CAA ACA CCT GAA GAC ATT TTG GAA	276
Ser Tyr Ala Lys Leu Val Gln Asn Gln Thr Pro Glu Asp Ile Leu Glu	
15 20 25 30	
GAG TTT CCT GAA ATT ATT CCA TTA CAA CAA AGA CCT AAT ACC CGA TCT	324
Glu Phe Pro Glu Ile Ile Pro Leu Gln Gln Arg Pro Asn Thr Arg Ser	
35 40 45	
AGT GAG ACG TCA AAT GAC GAA AGC GGA GAA ACA TGT TTT TCT GGT CAT	372
Ser Glu Thr Ser Asn Asp Glu Ser Gly Glu Thr Cys Phe Ser Gly His	
50 55 60	
GAT GAG GAG CAA ATT AAG TTA ATG AAT GAA AAT TGT ATT GTT TTG GAT	420
Asp Glu Glu Gln Ile Lys Leu Met Asn Glu Asn Cys Ile Val Leu Asp	
65 70 75	
TGG GAC GAT AAT GCT ATT GGT GCC GGT ACC AAG AAA GTT TGT CAT TTA	468
Trp Asp Asp Asn Ala Ile Gly Ala Gly Thr Lys Lys Val Cys His Leu	
80 85 90	
ATG GAA AAT ATT GAA AAG GGT TTA CTA CAT CGT GCA TTC TCC GTC TTT	516
Met Glu Asn Ile Glu Lys Gly Leu Leu His Arg Ala Phe Ser Val Phe	
95 100 105 110	
ATT TTC AAT GAA CAA GGT GAA TTA CTT TTA CAA CAA AGA GCC ACT GAA	564
Ile Phe Asn Glu Gln Gly Glu Leu Leu Gln Gln Arg Ala Thr Glu	
115 120 125	

AAA ATA ACT TTC CCT GAT CTT TGG ACT AAC ACA TGC TGC TCT CAT CCA Lys Ile Thr Phe Pro Asp Leu Trp Thr Asn Thr Cys Cys Ser His Pro 130 135 140	612
CTA TGT ATT GAT GAC GAA TTA GGT TTG AAG GGT AAG CTA GAC GAT AAG Leu Cys Ile Asp Asp Glu Leu Gly Leu Lys Gly Lys Leu Asp Asp Lys 145 150 155	660
ATT AAG GGC GCT ATT ACT GCG GCG GTG AGA AAA CTA GAT CAT GAA TTA Ile Lys Gly Ala Ile Thr Ala Ala Val Arg Lys Leu Asp His Glu Leu 160 165 170	708
GGT ATT CCA GAA GAT GAA ACT AAG ACA AGG GGT AAG TTT CAC TTT TTA Gly Ile Pro Glu Asp Glu Thr Lys Thr Arg Gly Lys Phe His Phe Leu 175 180 185 190	756
AAC AGA ATC CAT TAC ATG GCA CCA AGC AAT GAA CCA TGG GGT GAA CAT Asn Arg Ile His Tyr Met Ala Pro Ser Asn Glu Pro Trp Gly Glu His 195 200 205	804
GAA ATT GAT TAC ATC CTA TTT TAT AAG ATC AAC GCT AAA GAA AAC TTG Glu Ile Asp Tyr Ile Leu Phe Tyr Lys Ile Asn Ala Lys Glu Asn Leu 210 215 220	852
ACT GTC AAC CCA AAC GTC AAT GAA GTT AGA GAC TTC AAA TGG GTT TCA Thr Val Asn Pro Asn Val Asn GAA Val Arg Asp Phe Lys Trp Val Ser 225 230 235	900
CCA AAT GAT TTG AAA ACT ATG TTT GCT GAC CCA AGT TAC AAG TTT ACG Pro Asn Asp Leu Lys Thr Met Phe Ala Asp Pro Ser Tyr Lys Phe Thr 240 245 250	948
CCT TGG TTT AAG ATT ATT TGC GAG AAT TAC TTA TTC AAC TGG TGG GAG Pro Trp Phe Lys Ile Ile Cys Glu Asn Tyr Leu Phe Asn Trp Trp Glu 255 260 265 270	996
CAA TTA GAT GAC CTT TCT GAA GTG GAA AAT GAC AGG CAA ATT CAT AGA Gln Leu Asp Asp Leu Ser Glu Val Glu Asn Asp Arg Gln Ile His Arg 275 280 285	1044
ATG CTA TAACAACG Met Leu	1058

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TCGATGGGGG TTGCCTTTCT TTTTCGG

27

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "Primer"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGTTGTTA TAGCATTCTA TGAATTGCC

30